

Mapper des données sur des graphiques

Commençons

Vous ferez tout votre travail R dans R Markdown cette fois (et à partir de maintenant). Vous devez utiliser un projet RStudio pour bien organiser vos fichiers.

Vous devrez télécharger trois fichiers CSV et les placer quelque part sur votre ordinateur, de préférence dans un dossier nommé “data” dans votre dossier de projet :

- The_Fellowship_Of_The_Ring.csv
- The_Two_Towers.csv
- The_Return_Of_The_King.csv

Pour vous aider, j’ai créé un squelette de fichier R Markdown avec un modèle pour cet exercice. Téléchargez-le ici et incluez-le dans votre projet :

- 03-exercise.Rmd

Au final, la structure de votre répertoire de projet devrait ressembler à ceci :

```
your-project-name\  
  03-exercise.Rmd  
your-project-name.Rproj  
data\  
  The_Fellowship_Of_The_Ring.csv  
  The_Two_Towers.csv  
  The_Return_Of_The_King.csv
```

Tâche 1: Reflection

Écrivez votre réflexion pour les lectures de ce chapitre. Voici quelques questions pour vous aider:

- Pourquoi est-il important de visualiser les variables et les données ?
- Que signifie mapper des données pour représenter graphiquement l'esthétique ?
- Quelles données ont été mappées à quelle esthétique dans la vidéo de Rosling ?

Tâche 2: le Seigneur des Anneaux

Répondre aux questions suivantes:

1. Utilisez `group_by()` et `summarize()` sur les données `lotr` pour trouver le nombre total de mots prononcés par race. *
Ne vous inquiétez pas du graphique pour l'instant *. Combien de mots les hobbits masculins ont-ils prononcés dans les films ?
2. Utilisez `group_by()` et `summarize()` pour répondre à ces questions *avec des graphiques à barres* (`geom_col()`)
3. Est-ce qu'une certaine race domine toute la trilogie ? (indice : regrouper par "Race")
4. Est-ce qu'un certain sexe domine un film ? (bien sûr que c'est le cas, mais quand même, représentez-le graphiquement)
(Astuce : groupez à la fois par `Gender` et `Film`.) Expérimentez avec le remplissage par `Gender` ou `Film` et le facettage par `Gender` ou `Film`.
5. La race dominante diffère-t-elle d'un film à l'autre ? (Astuce : regrouper à la fois par `Race` et `Film`.) Expérimentez avec le remplissage par `Race` ou `Film` et le facettage par `Race` ou `Film`.
6. Créez un graphique qui visualise simultanément le nombre de mots prononcés par race, sexe et film. Utilisez le jeu de données "lotr" complet. Vous n'avez pas besoin de créer un nouvel ensemble de données résumées (avec `group_by(Race, Gender, Film)`) car les données d'origine ont déjà une ligne pour chacune d'entre elles (vous pourriez créer un ensemble de données résumées, mais il serait identique à la version complète).

7. Vous devez afficher `Race` , `Gender` et `Film` en même temps, mais vous n’avez que deux esthétiques possibles (`x` et `fill`), vous devrez donc également faire face à la troisième. Jouez avec différentes combinaisons (par exemple, essayez `x = Race` , puis `x = Film`) jusqu’à ce que vous en trouviez une qui raconte l’histoire la plus claire. Pour vous amuser, ajoutez un `labs()` pour ajouter un titre, un sous-titre et une légende.

Vous devrez insérer vos propres morceaux de code là où c’est nécessaire. Plutôt que de les taper à la main (c’est fastidieux et vous pourriez mal compter le nombre de backticks !), utilisez le bouton “Insérer” en haut de la fenêtre d’édition.

Tout remettre en place

Lorsque vous avez terminé, cliquez sur le bouton “Knit” en haut de la fenêtre d’édition et créez une version HTML ou Word (ou PDF si vous avez [installé tinytex](#)) de votre document.